APPENDIX O TOXICOLOGY OVERVIEW

O.1 BACKGROUND

Toxicology is the study of the adverse effects of chemical agents on biologic systems. It has been a discrete science since the mid-1970's with subdivisions for food, clinical, industrial, and environmental toxicology, but the Society of Toxicology has been in existence since 1961. The origins of the study of toxicology go back to ancient times.

The ***science*** of toxicology is observation and data-gathering. The ***art*** of toxicology is *predicting* the dose at which adverse effects will occur in a population outside the laboratory. Dr. A.J. Lehman is quoted (Casarett and Doull, 1986) as saying, "Anyone can become a toxicologist in two easy lessons, each of which takes ten years."

Philippus Aureolus Theophrastus Bombastus von Hohenheim-Paracelsus, who lived during 1493 to 1541, identified that

> *"All substances are poisons; there is none which is not a poison.*
> *The right dose differentiates a poison and a remedy."*

The corollary to Paracelsus' statement is that there is no chemical that cannot be used safely by limiting the dose or exposure. Similarly, a tenet of risk assessment/management is that where there is no exposure, there is no risk.

Industrial toxicology pertains to exposures in an industrial/occupational setting where exposure concentrations are often higher than those found in environmental releases. Industrial spills are sudden releases of relatively high concentrations of exposure agents compared to exposure concentrations in normal industrial operations, which themselves are usually greater than those found in investigations of historical environmental releases.

O.1.1 The Risk Assessment Paradigm

Toxicity data provide the basis for characterizing exposures for the significance of the exposure, and, in turn, the release. Toxicology and toxicity assessment are consistent with the NRC 1983 report, *Risk Assessment in the Federal Government*, establishing the risk assessment paradigm (model), shown in Figure O.1.1.

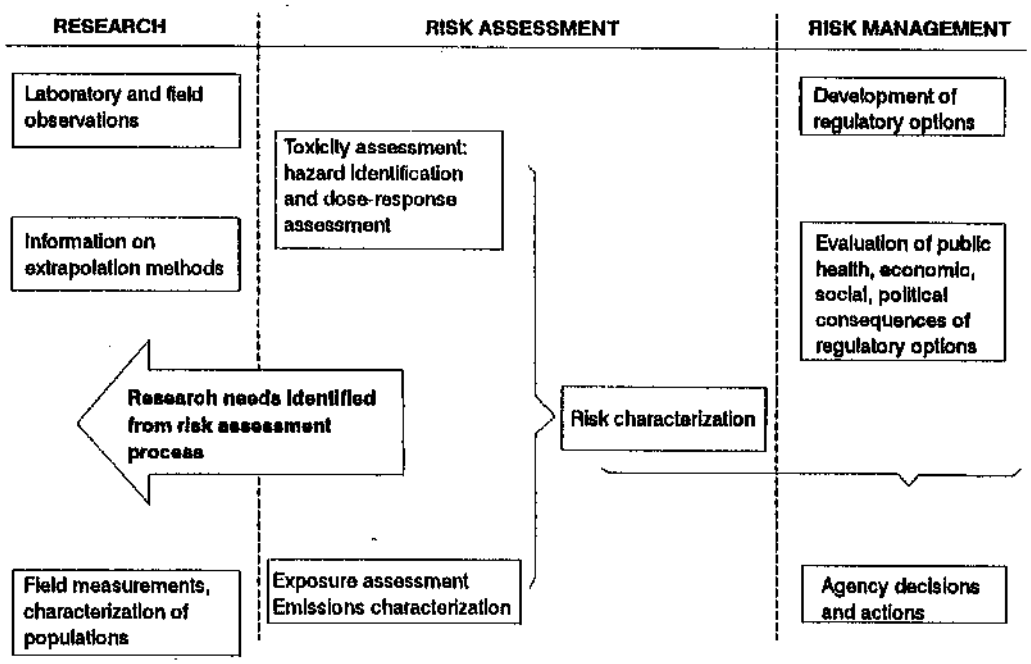


Figure O.1.1 NAS/NRC Risk Assessment/Management Paradigm (from NRC, 1983).

Implementation of the NRC paradigm, as in EPA's *Risk Assessment Guidance for Superfund* (1989a), has resulted in the model, shown in Figure O.1.2, often associated with environmental toxicological risk assessment.

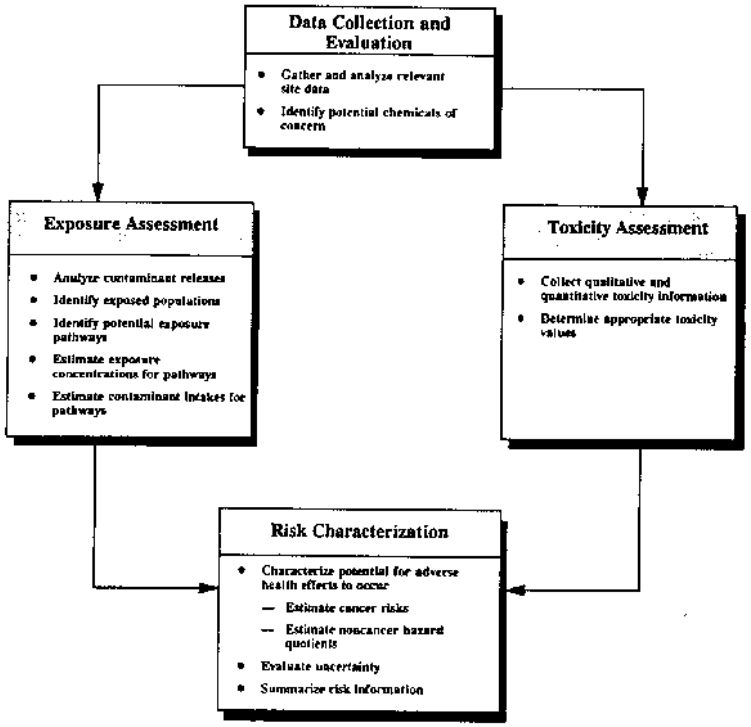


Figure O.1.2 EPA Risk Assessment For Superfund Paradigm.

The model indicates that information about hazardous chemical properties and site characterization, called **Data Collection and Evaluation** in Figure O.2, is evaluated to construct an **Exposure Assessment** consisting of an exposure scenario, a receptor population survey, and construction of an exposure dose. The exposure dose is compared to toxicity reference information from a **Toxicity Assessment** to prepare a **Risk Characterization** for risk assessment and subsequent risk management.

Obviously, the exposure dose calculated for the exposure assessment must be compatible with the dose cited from toxicity experiments that are part of the laboratory database in Figure O.1. A few words about “dose,” “dose rate,” “intake,” and “intake rate” are necessary to appreciate how the proper comparison of exposure assessment and toxicity assessment should be conducted.

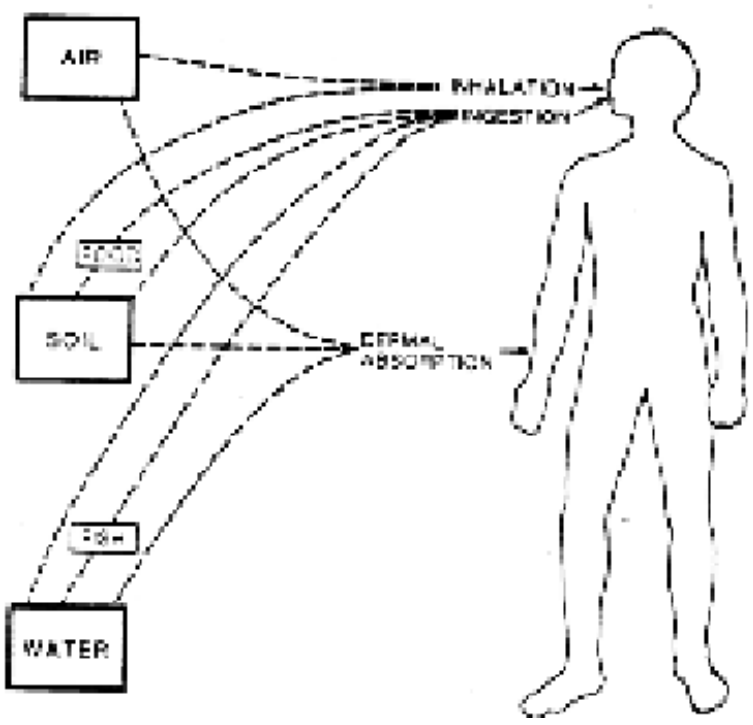
O.2 TOXICITY DOSE UNITS

The common term for the amount of chemical agent taken in to a receptor in an exposure is “dose”. For environmental exposure, the data are still growing but are far from complete describing how much of a chemical taken in is actually absorbed and linked to exposure effects. Therefore, environmental exposures are commonly based on the *administered* dose, rather than an *absorbed* dose. This carries with it the assumption that chemical concentrations presented to a receptor are ingested, or inhaled 100 percent. In cases where there is enough information to estimate respiratory exhalation of volatile COCs, data are provided in this guidance and can be used, with justification, as part of a Tier 2 evaluation. Incidental soil ingestion is assumed to be an administered dose without correction for absorption, just as in the clinical toxicity studies upon which the toxicity standards are based.

The amount of an agent of exposure entering the receptor is usually expressed as milligrams (mg) of the chemical (COC). The dose (or intake) is expressed as mg per kilogram of body weight. A dose or intake typically applies to acute, single exposures or dose regimens of relatively short duration. Exposures over a long period of time, *e.g.* months to years, are referenced to a rate, such as *mg/kg-day*, and are called a dose rate or intake rate (may also be found as mg/kg/day). This applies to chronic, continuing exposures, as in environmental exposure scenarios. It makes sense, then, to cite toxicity standards in the same units, *mg/kg-day*.

O.3 EXPOSURE ROUTES

Figure O.3 depicts the potential exposure routes for human exposure.



Not all of the exposure routes depicted are included in a Tier 1 risk-screening evaluation. For instance, deposition of airborne contaminants on skin followed by dermal contact absorption is small compared to potential exposure via incidental ingestion of soil, small enough that the pathway is not included, even in the conservative Tier 1 evaluation.

Unless there are site-specific features that include ingestion of crops grown on the site or fish raised in contaminated water on the site, these pathways are not included in the risk-based evaluation for UST sites. Given the usual configuration of a UST site, it would be unusual that crops or fish would be raised on site in Arizona and ingested.

Inhaled particulate, as much as 50% of total exposure (EPA, 1985), is deposited in the upper respiratory passages, where it is cleared from the lungs by ciliary-mucous processing followed by ingestion exposure. Usually, contribution to ingestion exposure by this process is small compared to the amount assumed for incidental ingestion of soil, and an exposure route for ingestion exposure via inhalation is not considered in the risk-based screening.

- **Inhalation exposure** involves entry into the body via the epithelial layer of the lungs to the blood to be carried throughout the body. Solubility of contaminants in lung tissue is important to evaluating exposure, but for risk-based evaluation, the assumption is made that, without substantiating evidence to the contrary, an exposure concentration encountered is an exposure concentration absorbed, *i.e.*, administered dose is assumed to be the same as absorbed dose.

- **Ingestion exposure** involves absorption of ingested contaminants through alveolar tissue in the digestive system to the blood to be carried throughout the body. At the relatively low levels of environmental exposure, there is a low likelihood of contaminant interaction, *i.e.*, synergisms-positive or negative, so the practice in risk-based evalution of adding HQ's is a good first estimate of total exposure. As with inhalation exposure, unless there is evidence supporting a particular absorption factor, the exposure point concentrations are assumed to be absorbed completely. This is consistent with the toxicity standard development process where exposure is almost always based on administered dose. Comparison to toxicity standards for risk-based evaluation needs to be on a common basis. IRIS (EPA, 2000) is set up on administered dose, but RAGS (EPA, 1989a) provides procedures for correcting toxicity factors for the use of absorption factors where justified.

- **Dermal contact absorption exposure** involves migration of contaminants through the lipid tissue of the skin to the blood where they are carried throughout the body. Not all chemicals are effective at penetrating the skin. Metals, for instance, migrate through the hair follicles and the effectiveness is generalized as 1% based on the skin area occupied by the follicles (Casarett and Doull, 1986). Other chemicals have varying degrees of effectiveness at crossing the skin barrier, and they have been characterized through tabularizations of chemical-specific permeability coefficients, K_p's (EPA, 1992). K_p values are incorporated in the risk-based screening concentrations of tier 1 and may be modified, with justification, in Tiers 2 and 3. For UST sites where BTEX and TPH are the primary constituents of a release, dermal contact absorption exposure via soil is about 8 -10 percent of the total estimated exposure for risk-based evaluation.

O.4 DEVELOPMENT OF TOXICITY STANDARDS

Toxicity standards are developed through a mulitiplicity of toxicity tests.

O.4.1 Clinical Toxicity Test Design

The elements in setting up a clinical study to develop a toxicity standard are represented in Figure O.4.1.

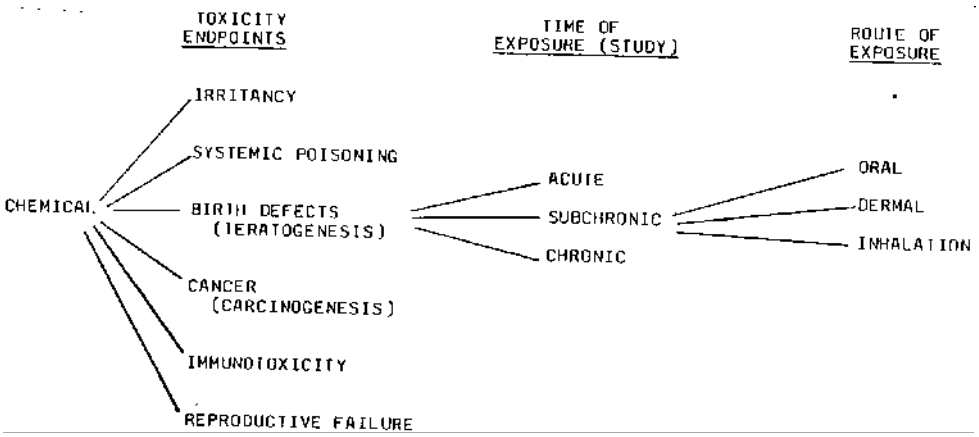


Figure O.4.1 Multiplicities in Clinical Toxicological Studies.

Considering environmental exposure and Figure O.4.1, the toxicity endpoints for a particular chemical would be either carcinogenesis (benzene, for example) or systemic poisoning (benzene, non-cancer endpoint, or toluene, a non-carcinogen, as examples). Either of these endpoints requires a clinical study costing $100,000 to $1,000,000 and running as much as two years in duration when rats are the test species. The duration of the toxicity study must be geared to type of exposure: acute (hours to a couple weeks), subchronic (usually less than 7 years in human exposure time), and chronic exposure (whole-life exposure). Environmental exposure, especially for risk-based evaluation, is considered chronic exposure and is further subdivided into oral (ingestion), dermal, or inhalation exposure routes, all requiring toxicity factors for risk-based evaluation. Toxicity standards for a single chemical in environmental exposure causing systemic health effects on a chronic basis via oral, dermal, or inhalation exposure routes, could require 3 separate clinical trials costing several million dollars. With the current state of knowledge, toxicity standards are being compiled for oral and inhalation exposure pathways. The oral toxicity standard is used to approximate the dermal toxicity standard in risk-based evaluation.

O.4.2 Dose-Response Relationships

The construction of the simple multiplicities in Figure O.4.1 is based on understanding of dose-response relationships. An alternate re-statement of Paracelsus' statement, cited earlier, is that the distinction between the therapeutic and toxic properties of an agent is distinguishable by the dose. Examples of these relationships are shown in Figure O.4.2.

In graph A, there is no response, good or bad, regardless of the dose. In graph B, increasing the dose reaches a threshold dose level where an

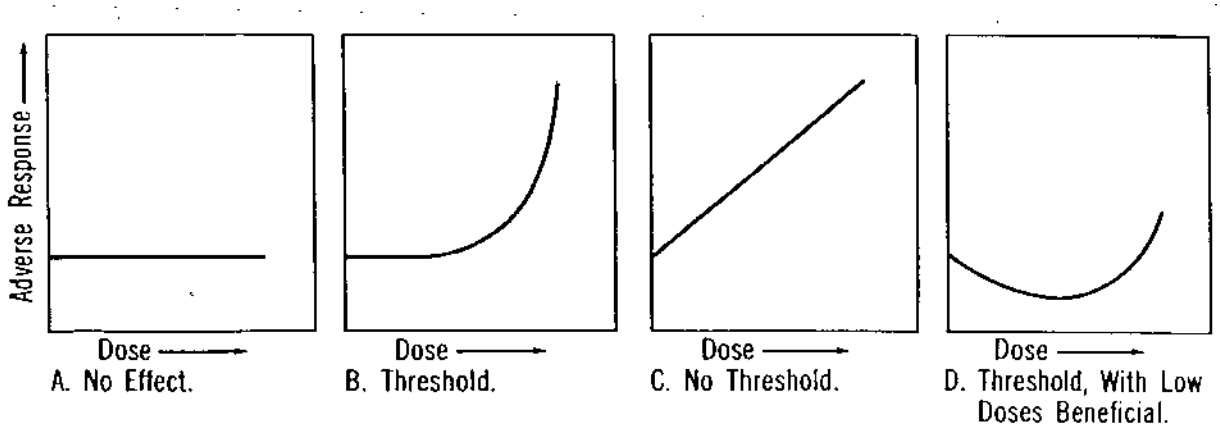


Figure O.4.2 Dose-Response Types.

adverse effect arises. In graph C, there is no threshold, but the adverse effect increases steadily with dose. In graph D, low doses produce a reduction in the adverse effect, a therapeutic effect, followed by increasingly adverse effect with higher doses. This kind of effect has been called a hormetic (or "U-shaped" curve) effect. The phenomenon of *hormesis* will be critical to our understanding of toxicity in risk-based evaluation in the years to come because some low-level, environmental exposures that have been considered detrimental in the past, are being re-considered in the light of hormetic effects. Reconsideration of the toxic effects of lead as both a potential carcinogen and as a systemic toxicant is an example. [For more information, consult the Belle Newsletter, www.BELLEonline.com.] For purposes of UST risk-based evaluation, carcinogens continue to be evaluated as if they behave as in graph C; there is no theshold, and minute exposures have a probability of producing an adverse cancer effect. Non-carcinogens are considered to be systemic toxicants behaving as in graph B; there is a threshold below which exposure over a lifetime by even sensitive sub-populations of the general population are protected.

O.4.3 Chronic Toxicity Factors

Dose-response information, as in Figure O.4.2, is converted into toxicity standards that can be used for risk-based evaluation.

O.4.3.1 Non-Cancer Effects -- The Reference Dose (RfD)

The reference dose is an estimate with uncertainty spanning perhaps an order of magnitude or greater) of a daily exposure level for the human population, including sensitive sub-populations, that is likely to be without an appreciable risk of deleterious effects during a lifetime (EPA, 1989a, page 7-5). For the determination of a Reference Dose (oral, for instance) for a non-cancer health effect, as in Figure O.4.3.1, the clinical experiment is set up by first determining a maximum tolerable dose (MTD), the maximum dose that the test organisms (animals, nominally rats for this illustration) can tolerate over the time of the experiment without any of the animals dying.

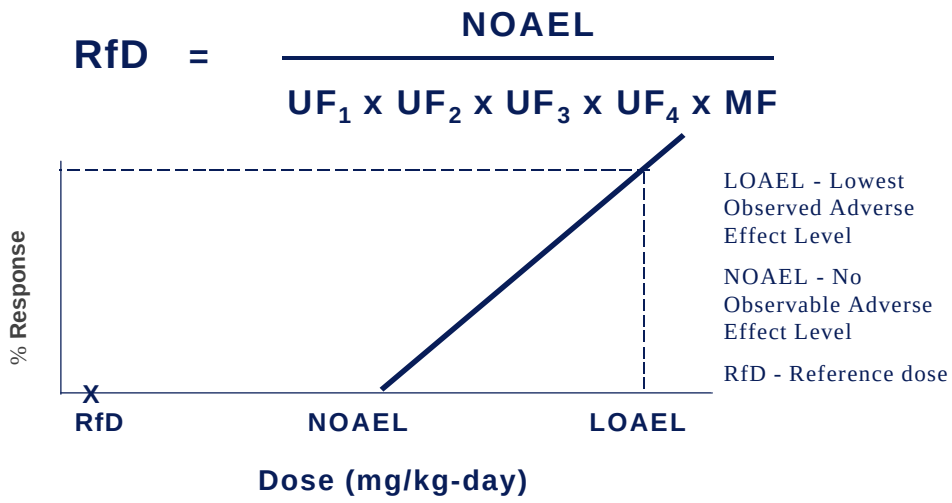


Figure O.4.3.1 Dose-Response Relationship for Non-cancer effects and Development of a Reference Dose.

Since chronic (whole-life) exposure is of interest, a two-year period, the life span of rats is usually used. Several lower doses than the MTD are chosen, and groups of test animals (usually 10 in a group) are administered the dose via the exposure route, oral in this case, of interest for two years as a whole-life exposure. A control group is maintained without exposure for comparison. [If chemicals are administered mixed in a syrup of safflower oil or corn oil, the control

group is fed only the oil.] The percent response in each group of rats for a particular dose, compared to the response in the control group, becomes a data point. A collection of data points at the comparatively high doses of the clinical experiment would be located to the upper right in Figure O.4.3.1, if they were shown.

The dose response points of interest for purposes of establishing an RfD are found at lower doses, beginning with the lowest observed adverse effect level (LOAEL). The LOAEL dose is not a good candidate as a reference dose for the following reasons:

- The LOAEL (or NOAEL) as single doses do not account for variation in the general population and do not provide protection for sensitive sub-populations, *e.g.*, elderly, children;
- The LOAEL must be extrapolated from animals to humans. Interspecies variability between humans and other mammals is included in this factor;
- There is uncertainty when a subchronic clinic study is used instead of a chronic study in determining a LOAEL; and
- Developing an RfD from a NOAEL (estimating a NOAEL) involves uncertainty in exptrapolating from the LOAEL to the NOAEL.

To account for these uncertainties, the LOAEL is divided by uncertainty factors, UFs, to address the four key uncertainties, thereby providing a margin of safety in the RfD for the general public. Because the UFs have a value of 10, an additional factor is sometimes used to incorporate additional knowledge and decreased uncertainty about one of the four areas of uncertainty. Additional uncertainties are addressed by applying a Modifying Factor, MF, along with the UFs. The MF is intended to reflect a qualitative professional assessment of additional uncertainties in the critical study and in the entire data based for the chemical not addressed by the preceeding uncertainty factors. By convention, values of 10, 7 or 3 are used as MFs, depending on the extent of the additional knowledge. When there is no additional information, MF has a value of 1. [The MF is set less than one for a small number of substances to account for nutritional essentiality.]

The result of these evaluations, conducted by expert toxicologists and review panels, are the chemical-specific reference doses listed in EPA's Integrated Risk Information System (IRIS; URL: *http://www.epa.gov/ngispgm3/iris/index.html*) to be used in risk-

based evaluations for UST facilities.

O.4.3.2 Cancer Effects -- The Carcinogenic Slope Factor (SF)

The clinical studies supporting toxicity standards for cancer effects are similar to those for non-cancer effects. The toxicity endpoint of interest is the formation of malignant tissue or tumors. The study design includes choice of administered doses that will provide information on the variation of tumor formation with dose. First, the study design is based on information that the chemical is likely to be a carcingen. That determination comes from a review of human and animal studies for **wieght-of-evidence**. The EPA weight-of-evidence classification system is listed as shown in the following table.

EPA's Weight-of-Evidence Classification System for Carcinogenicity*

Group	Description
A	Human carcinogen
B1 or B2	Probably human carcinogen
	B1 indicates that limited human data are available. B2 indicates sufficient evidence in animals and inadequate or no evidence in humans
C	Possible human carcinogen
D	Not classifiable as to human carcinogencity
E	Evidence of noncarcinogenicity for humans

*EPA, 1989a, page 7-11.

The weight-of-evidence of the data for a particular chemical is used to classify those carcinogens (and environmental cancer risk) in a UST risk assessment. Typically, benzene is a group A carcinogen and is the driver for risk assessment of a relatively fresh gasoline release. For other petroleum releases, PAHs might be involved. Some PAHs are non-carcinogens, but several have carcinogenic effects. The question of whether the risk for PAHs and benzene should be summed in a risk assessment is based on the weight-of-evidence classification. In general, risks from weight-of-evidence classes A and B (and C where slope factors exist for those

chemicals) should be summed. In every case, the weight-of-evidence group should be listed for each carcinogen in a risk assessment evaluation.

A graph of typical results from clinical animal studies is shown in Figure O.4.3.2. The data points, obtained at relatively high dose are shown about the least-squares line that describes the non-linear relationship of the data while minimizing the error between the data points and the line.

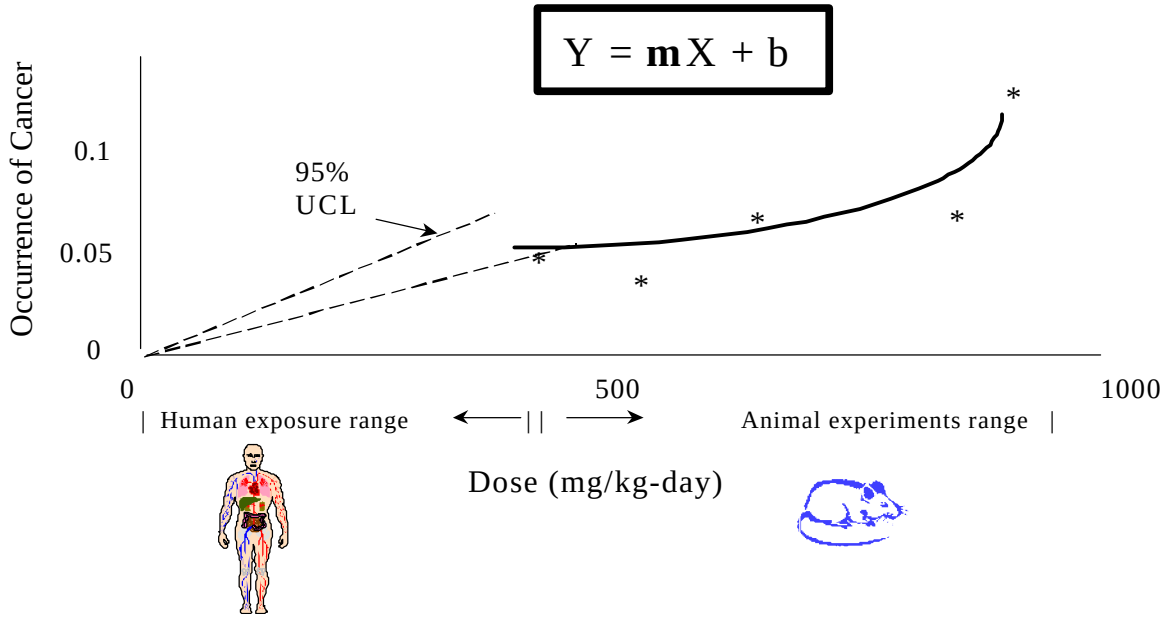


Figure O.4.3.2 Development of Slope Factor from Cancer Dose-Response Data

Environmental exposures to humans occurs at doses lower than the range of clinical animal studies. The challenge is how to express dose-response from the lowest measured data point to zero. Past theory has been that there is no threshold for cancer effects; one carcinogen molecule breaking the chain of one DNA strand will set a run-away, cancerous replication in motion that results in a tumor. Today we know that this theory does not hold true. In fact, it is estimated that repairative mechanisms in the human body fight off as many as 10,000 cancer-like assaults at the cellular level each day. EPA's proposed revisions to the cancer assessment guidelines in 1996 [http://www.epa.gov/neca/cancer.html] have not yet been

promulgated, and the methodology in use since the mid-1980's continues today.

Figure O.4.3.2 shows that, just as with the graph for non-cancer clinical studies, there is a data point for the percent of cancer responses compared to the minimum dose investigated. Hower, unlike with non-carcinogens, the data are evaluated as occurrence per unit dose. Where non-cancer effects are based on a safe threshold dose, cancer effects are assumed to occur from the minimum data point down to zero dose, zero response. The past and current practice has been to span the regeime of dose-response from the lowest data point to zero dose-zero response using a straight line (EPA, 1987). There is controversy in using this Linear-No-Threshold Model (LNT model) because research indicates that carcinogenesis usually occurs by a multi-step process that is unlikely to fit a single-line model (*cf.*, Kreeger, 1996). However, the proposed revisions to the cancer guidelines (EPA, 1996) have not be finalized, and the use of the LNT model is current practice.

In Figure O.4.3.2, a straight line is drawn from the data point at lowest dose to the origin of zero-dose/zero cancer occurrence, according to the LNT model. The straight-line graph makes possible the estimation of cancer occurrence at very low dose (and risk), characteristic of environmental exposures. The popular sentiment of protection of the public at a risk of one-in-one-million (10^{-6} risk) is enfranchised in EPA's National Contingency Plan (EPA, 1990). In fact, the upper bound of acceptable exposure is a risk on one-in-ten-thousand (10^{-4} incremental lifetime cancer risk), and 10^{-6} risk is designated as a *point of departure*. One important use of the *point of departure* is as a parameter for calculation of risk-based concentrations, called Preliminary Remediation Goals (PRGs) by EPA (EPA, 1991b; EPA1999b). Many of the equations in EPA, 1991b and 1999b are described elsewhere in this document as the basis of the risk-based concentrations for evaluation of UST releases in Arizona. The significance of 10^{-6} and 10^{-4} risks as regulatory criteria are also described elsewhere in this document.

The straight-line dose-response relationship in Figure O.4.3.2 allows development of a toxicity standard based on the equation of the straight line, $\boldsymbol{y}=\mathbf{m}\boldsymbol{x}+\boldsymbol{b}$. Since the line passes through the origin (0,0), $\boldsymbol{b}$, the intercept on the vertical ($\boldsymbol{y}$) axis is zero. The relationship between the dose and response is the slope of the line, **m**, in units of $(\text{mg/kg-day})^{-1}$.

Not yet accounted for by the slope of the first straight line of Figure O.4.3.2 are the uncertainties identified previously in the development of the reference dose. The technique for carcinogens is slightly different, but is directed to the same purpose. The second line in Figure 6-10 represents the line corresponding to the 95% upper confidence limit of the slope of the first line. This slope is the slope factor (SF) that serves as the chemical-specific toxicity factor for carcinogens. It represents an upper 95th percent confidence limit on the probability of a response per unit intake of a chemical over a lifetime (EPA, 1989a, page 7-12). The slope factors with their respective weight-of-evidence groups and the source of the data must be tabulated in the corrective action risk assessment for closure.

O.4.3 Other Toxicity Standards

In general, the shorter the exposure time, the higher is the allowable dose. EPA's Office of Drinking Water has developed One- and Ten-day Health Advisories for drinking water intake (http://www.epa.gov/ost/drinking/standards). While they are nonregulatory guidance, they serve a useful purpose for site-specific conditions where concentrations are greater than risk-based drinking water PRGs, GPLs, or AWQLs. Short-term criteria can be useful in indicating a more immediate threat that should be addressed through an interim remedial action that will protect human health and allow a more systematic approach to corrective action and final closure of the site. Site-specific evaluation for corrective action planning should always include a screening to see if there are immediate threats that should be addressed before the planned corrective action. In such case, communication with the ADEQ project manager for the site must be maintained to ensure that site-specific changes are known to all the stakeholders.